Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

A COMPREHENSIVE APPROACH TO OPTIMIZING COMPLEX SYSTEMS

Zelda B. Zabinsky and Pariyakorn Maneekul[1]

[1]Dept. of Industrial and Systems Eng., University of Washington, Seattle, WA, USA

ABSTRACT

Optimization viewed broadly can aid in designing, analyzing, and operating complex systems, from strategic policy planning to last-mile distribution to optimal control of dynamical systems. The optimization model, including decision variables, objective functions, and constraints, requires performance metrics that are often evaluated via black-box simulations. We summarize algorithms that can address black-box noisy functions, mixed integer- and real-valued variables, and multiple objectives. Multiple models (e.g., Gaussian processes, neural networks, queueing networks) can be used in conjunction with a computationally expensive model (e.g., simulation) to predict performance and reduce overall computation. A key issue in solving an optimization model is to dynamically allocate computational effort to efficiently search for the global optimum. The dilemma of exploration vs exploitation vs estimation is evident in machine learning and global optimization. We discuss sampling distributions that provide insights into this balancing act, and how ideas in quantum optimization provide approaches to optimizing complex systems.

1 INTRODUCTION

Optimizing a complex system has many aspects that must be considered before selecting a computational approach. The first consideration is the overall goal. An example is to design a future system. In this situation, a design engineer may wish to evaluate many designs while being limited by available data and models to predict performance. Multiple models may be available to estimate performance metrics with different assumptions and limitations. The design engineer may be interested in evaluating trade-offs between multiple objectives using an approximate set of Pareto optimal solutions and information on the efficient frontier. A sensitivity analysis may be important to account for uncertainty and investigate what happens if the model changes.

Another example of an overall goal is to improve the performance of an existing system where data and models are available to predict future behavior. In this example, there may be a large number of decision variables to evaluate, and they may involve both integer- and real-valued variables. The timing of the data and decisions are important and may vary from weekly, daily, or even real-time control.

Achieving the overall goal relies on creating a model, or models, of performance of the complex system. Often a discrete-event simulation is used to model a system. Analytical models, such as a queueing network, may also be useful. A model such as a neural network may be constructed from data, as in machine learning. Surrogate modeling or meta-modeling is another approach to constructing a model of a complex system. Multi-fidelity modeling aids in understanding performance of a complex system and reducing the computational burden. For example, function evaluations of a computationally expensive model can be interspersed with a less expensive model to reduce the computational effort.

The specification of an optimization model, including decision variables, objective functions, and constraints, may be fluid as the goal is modified. For example, a constraint may be moved into the objective function or vice versa. Data may be updated frequently or dynamically accessed. It is important to recognize noise in the data and noise in the models of performance. The formulation of the optimization model plays a role in the computational effort needed to achieve the overall goal.

While classical optimization algorithms are tasked with determining a single optimal solution, many researchers are recognizing the value of a set-based approach where a set of nearly optimal (or ε-optimal) solutions can allow decision makers to account for aspects of the complex system that are not included in the models of performance. As an extension of set-based optimization, methods to determine a set of Pareto optimal (non-dominated) solutions to multiple objective formulations enables decision makers to understand trade-offs between performance metrics.

Figure 1 illustrates different ways to interpret what is desired from the optimization. Suppose the optimal solution cannot be implemented exactly. In panel (a), the objective function value at x_2 is better (less than) the objective function at x_1, however, if the implementation may vary over an interval, then the solution at x_1 has less variability and may be the preferred solution. This highlights that the "optimal" solution may not be the goal of a user. Panel (b) illustrates a similar idea between feasibility and optimality. The solution x^* is at the edge of feasibility and has the minimum objective function value, however, a slight deviation in implementation of x^* may result in an infeasible solution.

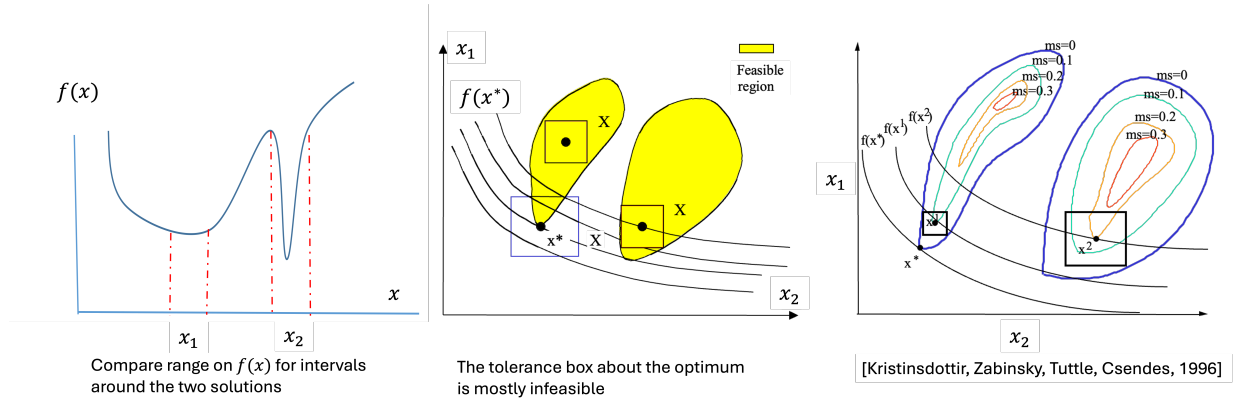

Figure 1: Illustration of tradeoffs between multiple objectives, exact solutions and slight deviations. Panels (a), (b), and (c) illustrate a tolerance box, showing variability of values when we cannot implement a solution exactly.

A motivating example of optimizing a complex system comes from engineering design of composite structures for aircraft fuselage. A team of University of Washington researchers and Boeing engineers worked on the optimal design of fuselage panels using composite laminates (Graesser et al. 1991; Graesser et al. 1993; Kristinsdottir et al. 1996; Kristinsdottir et al. 2001; Savic et al. 2001; Zabinsky et al. 2006). A finite element analysis of the panel served as a high-fidelity model and an analytical classical lamination theory model served as a low-fidelity model. Due to the nature of composite materials, the performance metrics were non-convex. The low-fidelity model was computationally much faster than the finite element analysis (seconds versus hours of CPU time), however, often an optimal design obtained using the low-fidelity model was infeasible when evaluated with the finite element analysis. Through manual adjustments of the limits of the design space, the team was able to provide designs that were useful to the design engineers.

In our experience, the design engineers were interested in trade-offs between multiple objectives such as margins of safety, manufacturing tolerances, weight, and cost. There were many types of margin of safety representing strength, strain, damage tolerance, and other metrics. A design is considered feasible if the minimum of the margins of safety is zero or greater. However, a larger margin of safety was desirable. Due to manufacturing considerations, an implementation of a specific design may vary within a "tolerance box". The cost was highly correlated with weight leading to three objectives: minimize weight, maximize the minimum margin of safety, and maximize the size of a feasible tolerance box.

Panel (c) of Figure 1 illustrates how a feasibility constraint and an implementation interval can be viewed as multiple objectives, taken from (Kristinsdottir et al. 1996).

Another manufacturing consideration was the level of discretization of fiber angles of plies in a composite laminate. While fiber angle of a ply mathematically is real-valued and may vary between $\pm 90^\circ$, the engineers were interested in considering different levels of discretization, such as restricting fiber angles to $\pm\{0, 15, 30, 45, 90\}$. We wanted an algorithm that could easily switch between allowable integer- and real-valued variables.

The need for solving a comprehensive optimization problem of a complex system was synergistic with analytical results on the power of sampling points in a design space according to adaptive random search with different distributions. Analyses of Pure Adaptive Search (PAS) (Zabinsky and Smith 1992; Zabinsky et al. 1995), Hesitant Adaptive Search (HAS) (Bulger and Wood 1998; Wood et al. 2001), and Annealing Adaptive Search (AAS) (Shen et al. 2007; Wood and Zabinsky 2002; Zabinsky 2003) provide insight into how a sampling distribution can balance exploration with exploitation and result in an efficient algorithm on black-box functions in a mixed integer/real variable domain.

These analyses illustrate the power of generating a point uniformly distributed in the current improving region. Zabinsky and Linz (2023) extend the analysis and provide insight into the role estimation plays in balancing exploration and exploitation. The analyses suggest that computational effort is better expended on discovering improving points than refining estimates of objective function values that may not be of interest during the progress of an adaptive search algorithm.

Although PAS is an idealized stochastic adaptive search algorithm and cannot be directly implemented, ideas in quantum optimization (Bulger et al. 2003; Baritompa et al. 2005) provide hope that computational methods may be able to implement optimization of large-scale, black-box, mixed integer/real variable problems efficiently in the future.

In this tutorial, we take a comprehensive view of optimizing complex systems. We discuss formulating an optimization model in Section 2. We provide an overview of global optimization in Section 3 including the use of multiple models and the role of estimation of noisy black-box functions. We summarize the power of sampling distributions in Section 4, and finish with a look to the future with quantum optimization in Section 5.

2 OPTIMIZATION MODEL

An optimization model can be formulated as

$$\begin{aligned} &\min_x \ f(x) \\ &\text{subject to } x \in S \end{aligned}$$

where the decision variables are denoted by a d-dimensional vector x. To be comprehensive, we allow the decision variables to include integer- and real-values. The domain S has d dimensions, with d_1 real-valued variables and d_2 integer-valued variables, where $d = d_1 + d_2$. Assume S is closed and bounded, and $S \subset \mathbb{R}^{d_1} \times \mathbb{Z}^{d_2}$. The objective function is typically denoted $f(x)$, $f : S \to \mathbb{R}$. The feasible region S can be determined by the intersection of constraints (e.g., $h_j(x) \geq 0$ for $j = 1, \ldots, J$) or by an oracle, that is, given a solution x, the model can return whether x is in S or not. We also consider the objective function to be a black-box determined by an oracle, that is, given a solution x, the oracle can return a value $f(x)$.

A multi-objective optimization model can be formulated with m multiple objective functions, as

$$\begin{aligned} &\min_x \ f_1(x), \ldots, f_m(x) \\ &\text{subject to } x \in S \end{aligned}$$

where each objective function maps S to the reals. When formulating a multi-objective optimization model, the goal may be to approximate the Pareto optimal set, that is, the set of non-dominated solutions. A point

$x \in S$ is Pareto optimal if there does not exist another point $y \in S$ such that $f_\ell(y) \leq f_\ell(x)$ for all $\ell = 1, \ldots, m$ and $f_\ell(y) < f_\ell(x)$ for at least one $\ell \in \{1, \ldots, m\}$.

It is common to assume the objective functions and constraints are deterministic and can be evaluated exactly, however the functions may need to be estimated using a noisy function. When the objective function $f(x)$ cannot be evaluated exactly, instead, we have a noisy evaluation available, that is, the performance at a design point $x \in S$ is given by $g : S \times \Omega \to \mathbb{R}$, where ω_x is a random element over a probability space denoted $(\Omega, \mathscr{A}, P)$. Then

$$f(x) = \mathbb{E}_\Omega[g(x, \omega_x)].$$

Depending on the decision maker's risk preferences and overall goals, the objective function can be reformulated by substituting the expectation with alternative risk measures, such as Value-at-Risk (VaR) or Conditional Value-at-Risk (CVaR) (Rockafellar and Uryasev 2000). Whereas the expected value measures a central tendency, VaR and CVar cover the tail-behavior of the distribution and specify values of a loss random variable associated with x at any specified probability level α, $\alpha \in (0,1)$. The objective function may be written as

$$f(x) = \text{VaR}_\alpha(x) \text{ or } f(x) = \text{CVaR}_\alpha(x).$$

In addition to the objective functions having noise, the constraints may also be stochastic. A common formulation is to involve the expected values in inequality constraints. This uncertainty can lead to risky or undesirable outcomes with low probability. A formulation that encompasses risk can be addressed with chance constraints, which specifies a certain tolerance for undesirable outcomes,

$$P(g(x, \omega_x) \geq 0) \geq \alpha,$$

where $\alpha \in (0,1)$ is the risk tolerance. Two main challenges of chance-constrained optimization are how to evaluate the probability of an undesirable solution and the non-convexity of the feasible set (Küçükyavuz and Jiang 2022). An advantage to using $\text{CVaR}_\alpha(x)$ is that it is a convex and coherent measure of risk (Uryasev and Rockafellar 2001). Additionally, CVaR is a natural upper bound for the VaR measure, making it a more conservative means to measure deviation from a given chance constraint. Many applications have a goal of limiting downside risk, such as portfolio optimization, power systems with renewable resources, health care systems, and supply chain optimization.

The goal of the optimization model may be to determine a single point that approximates the global optimum, or a set of solutions that approximate a level set, as in (Mason et al. 2022), a set of solutions that achieves a target quantile (e.g., best 10%), as in (Zabinsky and Huang 2020), or the Pareto optimal set. A benefit of providing a set of solutions enables decision makers to make trade-offs between simulated performance and other issues that may not enter into the model of performance. When there is noise in the performance measure, decision makers may be indifferent to small differences in the estimate of the objective function, and appreciate a set of solutions to further investigate.

3 GLOBAL OPTIMIZATION

Given an optimization model with a complete formulation, we turn to methods for providing solutions. There are many excellent papers that survey algorithms for solving black-box problems with noise. Before discussing them, we take a wholistic view of optimization.

In black-box optimization, the interpretation of a "local" solution versus a "global" solution depends on the neighborhood of a solution in the feasible region S. If the neighborhood is very large, that is, all of S, then a local solution *is* a global solution. The neighborhood is related to the algorithm. For example, in a Travelling Seller's Problem (TSP), if an algorithm is based on a one-city swap, then neighbors of a solution are those that can be obtained by swapping one city in the route (Notice gender neutral terminology with the use of the Latin ending "er."). If an algorithm is using a two-city swap, then neighbors can be reached by swapping two cities.

In Euclidean space, a neighborhood may be defined by a ball of a specific radius around a point. If the radius is very small, then neighbors to a point are very close, but if the radius is very large, as in Very Large Neighborhood Search (Ahuja et al. 2002), then every point is a neighbor to every other point. Algorithms based on a Markov chain Monte Carlo sampler, such as Hit-and-Run (Smith 1984; Zabinsky et al. 1993), make it possible to move to any point in the space in one step. In this case, every point in the entire domain is a neighbor to every other point in the domain although some neighbors may be "closer" or more easily reached probabilistically. The notion of "local" versus "global" may lose its meaning when approaching optimization of complex systems wholistically.

As Locatelli and Schoen (2021) note in their paper on (global) optimization:

> A final observation before the beginning of the paper: here and in the title we parenthesized the word (global). We will omit doing so in the paper, but we would like to observe that, although in the past the subject was considered somewhat exotic and off the main research streams, nowadays the richness of both theory as well as computational approaches gives to the subject a full recognition in the scientific community. We might then propose, with a slightly provocative style, to rename the whole subject simply as "optimization" – what else should we look for when optimizing, if not a global optimum? Locatelli and Schoen (2021)[p.1].

A challenge for optimization methodologies is that "*global optimization requires global information.*" Törn and Žilinskas (1989) show that deterministic algorithms which use only function values at sample points converge to the global optimum on all continuous functions if and only if they search a dense set. Stephens and Baritompa (1998) extend this and develop analogous results for stochastic algorithms. However, dense sampling is not the only way to obtain global information. For example, information on the structural form of the objective function, such as convexity, Lipschitz constant, and upper and lower bounds can provide ways to determine optima without sampling densely.

Statistical information, such as estimating confidence intervals or quantiles, is also a valid technique for gaining global information. De Haan (1981) shows that a random sample of points from a uniform distribution can be used to construct a confidence interval for the minimum using general extreme value theory. The Optimal Computing Budget Allocation (OCBA) methods for optimization under uncertainty use extreme value theory to provide statistical analyses that guide computation (Chen and Lee 2011).

3.1 Black-box Optimization

Black-box optimization is broadly applicable across many domains and has been studied in multiple scholarly fields under names including Derivative-free Optimization, Global Optimization, Bayesian Optimization, Sequential Experimental Design, and assorted variants of the multi-armed bandit problem. Methods include variations of simulated annealing (SA), genetic and evolutionary algorithms, partitioning methods, meta-modeling, particle swarm, and covariance matrix adaptation evolution strategy (CMA-ES) (Hansen and Ostermeier 2001), to name a few. Rios and Sahinidis (2013) review derivative-free optimization methods and Locatelli and Schoen (2021) survey global optimization methods. Bajaj et al. (2021) discusses black-box optimization methodologies and applications. Methods aimed at simulation-optimization are discussed in (Fu 2015; Gosavi 2015; Amaran et al. 2016; Shashaani 2024).

Several black-box optimization algorithms have been developed to accommodate a mixed integer/real feasible region. A review of algorithms on optimization problems with mixed integer/real variables is discussed in (Ploskas and Sahinidis 2022).

Algorithms originally designed for black-box optimization with a single objective have often been extended to allow for multiple objective functions. Since simulation can easily evaluate multiple performance measures of a system design, many simulation-optimization algorithms have been extended for optimizing complex multi-objective problems (Yoon and Bekker 2020). Algorithms for multi-objective problems include

evolutionary algorithms (Deb 2001; Sarker et al. 2002), particle swarm (Yang et al. 2011), interacting-particle (Mete and Zabinsky 2014), and weighted optimization using hierarchical bandits (Al-Dujaili and Suresh 2019). The OCBA method was adapted to solve multi-objective stochastic problems (Chen and Lee 2009; Li et al. 2017), and the Multi-Objective Convergent Optimization via Most-Promising-Area Stochastic Search (MO-COMPASS) was developed by Li et al. (2015). Partition-based algorithms have also been applied to multi-objective problems (Shi and Ólafsson 2009; Huang and Zabinsky 2014), as well as Kriging-based algorithms (Rojas-Gonzalez and Van Nieuwenhuyse 2020). The multi-objective genetic algorithm NSGA-II (Deb et al. 2002) is a widely used multi-objective algorithm (Yusoff et al. 2011; Verma et al. 2021).

3.2 Multiple Models

Often black-box optimization depends on the use of computationally expensive objective function evaluations. Consequently, a substantial body of research has been directed towards the development of methodologies to find good solutions under limited computational cost or time.

A widely used approach to reduce computation of expensive function evaluations is to use cheaper surrogate models, or meta-models (Barton 2020), as approximations. Popular surrogate models include Kriging, Gaussian process models, polynomial regression, moving least-squares and radial basis functions (Vu et al. 2017). Fitting the right surrogate model specification to discover and take advantage of the inherent structure of the objective function aids in a more accurate approximation, consequently a more efficient use of computational power.

Bayesian optimization uses a Gaussian process model to approximate the objective function and uses an acquisition function to determine the next sample point to evaluate with the expensive function. Popular examples of acquisition functions include expected improvement, knowledge gradient, and predictive entropy search. Significant computation arises from optimizing the acquisition function, particularly as the dimensionality of the problem increases (Wilson et al. 2018), posing a challenge for Bayesian optimization.

Multi-fidelity optimization typically refers to constructing separate high- and low-fidelity models based on knowledge of the complex system, where the high-fidelity model is assumed to be the more accurate model with a higher computational cost. Many engineering applications use a finite-element analysis with a fine-scale grid as the high-fidelity model and use a coarse-scale grid as the low-fidelity model. In our motivating example of design of composite structures, the high-fidelity model was a finite-element analysis and the low-fidelity model was an analytical model based on classical lamination theory. In (Morey et al. 2024), the high-fidelity model was a discrete-event simulation of a biomanufacturing system and the low-fidelity model was an analytical queueing Jackson network model. Methods to manage the use of both models (i.e., how many low-fidelity evaluations to perform between high-fidelity evaluations) are reviewed by Peherstorfer et al. (2018).

A major challenge with multi-fidelity modeling is capturing the complex relationships between the high- and low-fidelity models. Li, K. and F. Li (2024) discuss several approaches to align data and flow of information between the models to address this challenge. Zabinsky et al. (2019) utilizes a combination of additive modeling and partitioning techniques to build multi-fidelity models, where the additive model serves to define the correlation structure between the high- and low-fidelity observations. When the low-fidelity model demonstrates adequate accuracy, partitioning strategies are implemented to facilitate the exploration of the global optimum. However, if the accuracy is insufficient, evaluations of the high-fidelity function are acquired to refine the low-fidelity model.

Multi-fidelity modeling often operates under the assumption that the high-fidelity model serves as a “ground truth” or offers sufficient accuracy to represent the system of interest. However, in practical applications, the available high-fidelity model may not always fully capture the system’s behavior, leading to a "no ground truth" situation. In our experience of optimizing the design of composite structures, we realized that the finite element analysis and analytical solutions based on classical lamination theory exhibit inconsistent (and inaccurate) behavior near the boundary of the feasible region (Zabinsky et al.

2006). Another practical example presents in multi-fidelity topology optimization methodology for maritime environmental survey operations (Morey et al. 2021). The low-fidelity analytical models developed for this maritime operation scenario are able to accurately identify the intuitive maximum redundancy of coverage, whereas the high-fidelity simulations required modifications to represent accurate estimates of coverage.

Research on the use of multiple models when there is "no ground truth" is on-going, and developing measures of consistency between models to inform the optimization method. Xu et al. (2016) and Morey et al. (2024) propose new approaches to use consistency between models in the absence of a ground truth.

3.3 Estimation of Noisy Functions

The estimation of an objective function, e.g., estimating an expected value or CVaR using noisy function evaluations, adds another aspect to computational effort. Many methods rely on replications of a noisy function for estimation.

Another approach is a single-observation framework, i.e., the single observation search algorithm (SOSA), where only one replication is evaluated at any sampled point. SOSA was originally introduced for real-valued problems (Kiatsupaibul et al. 2018), and then extended to mixed integer- and real-valued variables (Kiatsupaibul et al. 2020). The left panel of Figure 2 illustrates estimating the objective function using a sample mean over R replications. The statistic obtained from the replications (e.g., mean, quantile, confidence interval) has convergence properties, following a frequentist statistic perspective, according to a central limit theorem. SOSA estimates the objective function of a sampled design point by averaging the single observations of the function values at different nearby points, see the right panel of Figure 2. Convergence of SOSA relies on the martingale property of the errors. Under mild regularity conditions, the optimal value estimate from the SOSA framework converges to the true optimal value with probability one for both continuous and mixed integer problems. Numerical experiments demonstrate the efficiency of the single observation framework on stochastic optimization problems, see Linz et al. (2017).

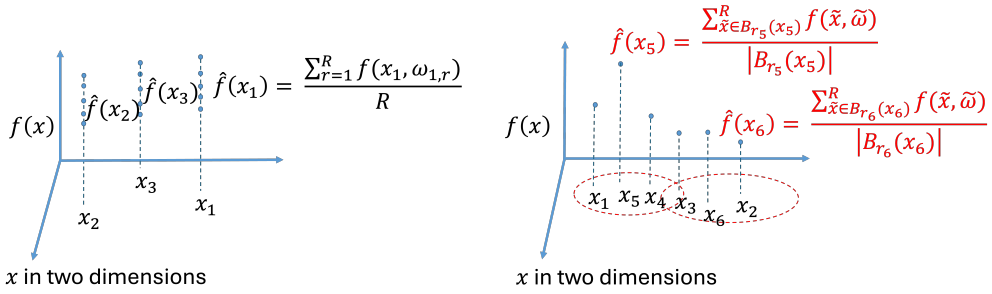

Figure 2: The left panel illustrates estimating an objective function value by averaging multiple observations of a noisy function value at the same x. The panel on the right estimates an objective function value by averaging single observations of the function values at different nearby points.

Independently created, the structure of SOSA inadvertently resembles the k-nearest neighbor algorithm in machine learning, making SOSA a machine learning algorithm for simulation optimization. However, SOSA employs averaging as its prediction function, whereas we recognize the power of basis expansion that is a generalization principle in machine learning. We propose incorporating quadratic basis functions into SOSA's prediction mechanism in the place of simple average. This enhancement with machine learning techniques aims to provide a more flexible and accurate representation of the objective function, leading to improved global optimum prediction.

3.4 Interplay between Optimization and Machine Learning

Algorithms for black-box optimization adaptively sample points in the feasible region and observe their corresponding objective function values. This information is used to predict the location of a global optimum and steer subsequent sampling towards a global optimal solution. On the other hand, a statistical method or machine learning algorithm takes sampled pairs of feature vectors (design points) and their corresponding responses (objective function values) and uses them to produce an inferred function to predict the responses of other unseen feature vectors. The two paradigms of prediction share many similarities. Both take sampled pairs of design points and their responses, and both use the information to predict a target response. A difference between these adaptive search algorithms and machine learning is the availability of design points and responses. For example, in engineering design, a response to a design point may involve running a computationally expensive simulation and evaluation may be sequential. A contrasting example is a recommender system with millions of design points and responses available, which is well-suited for a machine learning algorithm.

Due to its wide range of applications, machine learning has undergone a staggering development over the past decades and a large number of useful tools have been developed. Supervised learning is a type of machine learning that aims to predict the output (function value) of an input (feature vector or design point) based on an inferred function learned from a set of known input-output pairs. Some of the most well-known supervised learning algorithms include k-nearest neighbor (Devroye et al. 1994), linear regression, polynomial regression, the support vector machine (SVM) (Boser et al. 1992), the regression tree (Breiman et al. 2017) and neural networks (Goodfellow et al. 2016). The abstraction of these algorithms employs the concept of basis expansion (e.g., piecewise linear, polynomial, and radial basis functions) to achieve flexible representations of the target functions. Many black-box optimization algorithms for stochastic optimization employ these concepts, but we see an opportunity to go further in leveraging these methods.

3.5 Exploration, Exploitation, and Estimation

A crucial step in all of the algorithms for black-box optimization is how to strategically select the next point to evaluate. Whether a single model or multiple models are used, it is important to allocate the computational effort to balance exploration of the domain with exploitation of promising regions and estimation of noisy function evaluations.

Algorithms have different ways of balancing exploration, exploitation and estimation. One example of an algorithm with clearly distinguishable exploration and exploitation phases is a multi-start method. A multi-start method typically generates a point uniformly distributed over the domain (exploration phase) and then executes an algorithm such as a gradient descent around the neighborhood of a point (exploitation phase). For population-based algorithms, the "elite points" emphasize exploitation and other possibly random points stress exploration. In genetic algorithms, the crossover procedure tends to exploit points and the introduction of a mutation serves to explore. In simulated annealing, a high temperature parameter is effectively exploring, while a low value is exploitive. The acquisition function in Bayesian optimization can prioritize exploitation when it samples at the predicted location of the optimum, while sampling where the uncertainty of the Gaussian process model is large focuses on exploration. CMA-ES maintains a population of points and exploits promising points to update the mean of a multivariate Normal distribution, but also updates a covariance matrix and samples from the distribution to inject exploration into the algorithm. Partitioning is another means to balance exploration and exploitation, where focusing on a promising subregion is exploitation while keeping subregions with high variability of uncertainty of the objective function provides exploration. When estimating a function value with noisy evaluations, the inclusion of replications or single observation adds the computational effort for estimation into the balancing act.

We noted earlier that the neighborhood and algorithm are related. We also observe that algorithms induce an implicit sampling distribution. Analyses of the impact of sampling distributions on computational

effort can provide insight into how to guide an implicit sampling distribution through algorithmic actions to balance exploration, exploitation, and estimation.

4 GLOBAL OPTIMIZATION VIA SAMPLING DISTRIBUTIONS

To gain insight into the balancing act of exploration and exploitation, we can analyze the impact of sampling distributions on computational effort.

The sampling distribution of PAS, described in (Zabinsky 2003), is uniform on the set of improving points from the previous point. The idea of having a set of improving points may be interpreted as exploiting the objective function value of the best point sampled so far. Uniform sampling on the set of improving points could be considered exploring, in contrast to seeking the optimum in the level set. The finite-time analysis of PAS shows that, if one *could* sample uniformly from improving regions, then the expected number of such iterations to achieve a solution arbitrarily close to the global optimum with high probability increases at most linearly in dimension. We refer to this as the "linearity result."

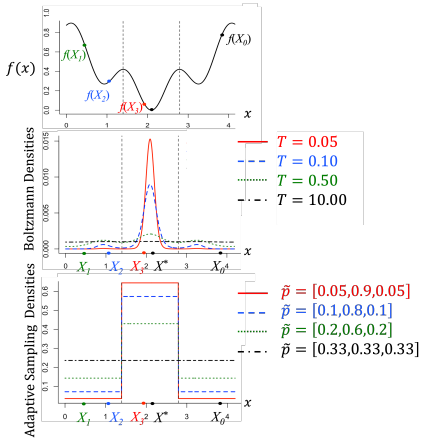

Figure 3: Sampling from a Boltzmann distribution parameterized by temperature induces the densities to focus more on the optimum as the temperature decreases. In BASSO, the adaptive subregion probability parameter $\tilde{p}$ impacts the sampling distribution.

While PAS is not directly implementable, the introduction of HAS (Bulger and Wood 1998; Wood et al. 2001) and AAS (Romeijn and Smith 1994; Shen et al. 2007; Wood and Zabinsky 2002) attempted to narrow the gap between theory and implementation. Annealing adaptive search is an idealized version of simulated annealing, where sampling from a sequence of Boltzmann distributions with decreasing temperature focuses the sampling distribution on improving level sets. Figure 3 illustrates how the Boltzmann distribution reflects the set of improving points as temperature decreases. With an appropriate cooling schedule, AAS can achieve the desired linearity result of PAS (Shen et al. 2007), although it is still impractical to sample from a Boltzmann distribution.

Many of the black-box optimization algorithms mentioned in Section 3.1 use partitioning and surrogate modeling as a means to search for a global minimum. Recently, an adaptive stochastic search algorithm called Branching Adaptive Surrogate Search Optimization (BASSO) (Maneekul, P. and Z. B. Zabinsky and G. Pedrielli. 2025) has been introduced that conceptualizes the use of branching and surrogate modeling. The bottom panel of Figure 3 illustrates how strategic partitioning can be used to steer the sampling distribution towards the improving level set of points. A finite time analysis of BASSO shows that the desired linearity result of PAS is achievable with specific assumptions and conditions.

These analyses provide insight into the exploration/exploitation trade-off. The sampling distribution must exploit the threshold of the best function value observed but must fully explore the associated level set of improving points.

Figure 5 illustrates the power of sampling from improving points with partitioning and uniform sampling. In this example, there are 400 grid points and the target level set of 10% quantile is outlined with a thick green contour. In the top panel, the domain is branched into four subregions and two points are sampled uniformly in each subregion. As in PBnB (Zabinsky and Huang 2020), the target quantile is estimated and confidence intervals on the estimate are shown with red dashed lines on the histogram in the top panel.

The next iteration branches each subregion creating 16 subregions, and points are uniformly sampled so that there are two points in each subregion. Using the confidence intervals on the estimated quantile, the PBnB analysis shows that the shaded subregions can be "pruned" with statistical confidence that they do not contain points in the target level set and are not worth expending further computational effort.

A third iteration branches only the current subregions and additional points are generated uniformly. Comparing the histograms of the newly sampled points in the three panels, we can see that by the third iteration, the sampling distribution is concentrated on a reasonable approximation of the target level set. This illustrates that uniform sampling on improving subregions can focus the overall sampling distribution and achieve efficient performance.

To consider the impact of estimation on the trade-offs between exploration and exploitation, Hesitant Adaptive Search with Estimation (HAS-E) (Zabinsky and Linz 2023) was introduced with a finite-time analysis of algorithm performance that combines estimation with a sampling distribution through the use of confidence intervals on the estimated function evaluations. An interpretation of HAS-E is that there is a tradeoff between sampling from a larger than needed level set (with loose upper confidence bound and fewer replications) and sampling from a more accurate estimate of the current level set (with tight upper confidence bound and more replications). This suggests that algorithms should use few replications as long as the estimation approaches the true function value as the algorithm approaches the global minimum.

While PAS, HAS, AAS, BASSO, and HAS-E are not directly implementable, the analyses provide insights into efficient features of a black-box optimization algorithm. The final insight is to not expend computation on regions with poor performance, and quickly focus the sampling distribution on improving regions. We now turn to ideas in quantum computing that may lead to a possible implementation of PAS.

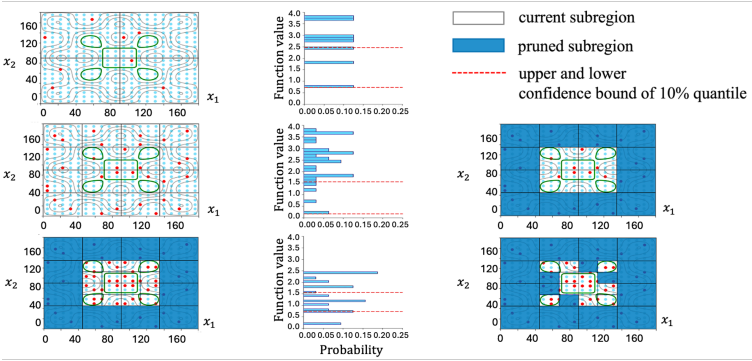

Figure 4: Approximating the 0.2-quantile level set (contour shown in green) on the centered sinusoidal function in two dimensions using partitioning algorithm with uniform sampling.

5 QUANTUM OPTIMIZATION

Whereas stochastic adaptive search algorithms have been attempting to approximate the ideal performance of PAS that is scalable in dimension (on average), quantum computing provides hope that this goal may be achievable in the future. Quantum computers inherently capture randomness, which is an important

characteristic of PAS. In contrast to conventional computing, where, given an input x an output $f(x)$ is returned, in a quantum computer, the output is a probability distribution with mean $f(x)$.

Quantum computing appears to be able to provide a natural implementation of sampling distributions that focus on improving level sets. Bulger et al. (2003) introduces Grover's adaptive search that provides a quantum implementation of PAS and HAS. The quantum annealing algorithm uses quantum tunneling to implement a quantum version of simulated annealing. A brief discussion of Grover's Search and quantum annealing follows.

5.1 Grover's Search

Grover's (Grover 1996) and Shor's (Shor 1994) algorithms are breakthroughs in quantum computing that have demonstrated its potential. Grover's algorithm, in particular, locates a single item in N items using $O(\sqrt{N})$ iterations of a Grover rotation, composed of a selective phase shift and a Grover operator. For finding m solutions out of N, the required number of Grover rotations is bounded by $\lceil(\pi/4)\sqrt{N/m}\rceil$. Grover's search algorithm has been applied to a discrete optimization problem, in particular, finding the minimum among an unsorted set of N different objects by Dürr, C. and P. Hoyer (1996). Their optimization implementation of Grover's search uses exponential searching (Boyer et al. 1998) by randomly selecting a possible solution, using its functional evaluation as the threshold in the selective phase shift operator, and applying a certain number of Grover rotations for each optimum search iteration.

Grover's adaptive search (GAS) is a framework that combines PAS with Grover's search algorithm (Bulger et al. 2003). GAS uses an adaptive search strategy to dynamically change the number of Grover rotations and demonstrates how GAS can be an implementation of PAS. It can also be viewed as a quantum-computational implementation of HAS. Grover Adaptive Search finds the optimum value of an objective function by using the best-known value from the previous run as a threshold. Setting a threshold for a new iteration from an earlier iteration in GAS is analogous to finding an improving level set in PAS. The adaptive oracle used in GAS recognizes all values above or below the current threshold (for maximize and minimize respectively), decreasing the size of the search space every iteration the threshold is updated, until an optimum is found. GAS then performs an amplitude amplification (inverting the amplitude of the current qubits state), which increases the probability of landing in the improving level set. If a better solution is found, the threshold is updated and a new iteration is started until the stopping criterion is met.

Baritompa et al. (2005) refines GAS where the number of Grover rotations for each iteration is determined by maximizing the benefit-cost ratio of the expected gain to the number of rotations. Bulger (2007) then addresses the application of Grover's algorithm with local search techniques where Grover's algorithm is used to locate the promising region that contains the global optimum solution, thereby combining the benefits of a multistart method with GAS. Liu and Koehler (2010) provided a different strategy to determine the benefit-cost ratio with Bayesian update. GAS has been extended to continuous optimization problems by discretizing the space using fixed-point representation (Protopopescu and Barhen 2005).

Morimoto et al. (2024) recently proposes Quantum Adaptive Distribution Search (QuADS), an extension of GAS using an adaptive multivariate normal distribution, mimicking CMA-ES for the initial state in quantum search. Numerical experiments conducted on QuADS demonstrate promising outcomes, highlighting the potential of quantum computing in tackling continuous optimization problems.

5.2 Quantum Annealing Algorithm

The quantum annealing algorithm (QAA) is an optimization algorithm that makes use of simulated quantum (rather than thermal) fluctuations and tunneling, thus providing a quantum-inspired version of simulated annealing. The property of quantum computing that allows the implementation of QAA is quantum tunneling. In sharp contrast to particles, quantum wave functions can tunnel through high potential barriers with significant probability, and this is formally known as quantum tunneling. The basic idea of QAA is to map the optimization problem to a physical system, such as a network of coupled qubits, and then

find the state of the system that corresponds to the minimum energy by gradually evolving the system to its ground state, using a process similar to the annealing process in macroscopic physics. Intuitively, this can be viewed as global evolution and superposition of quantum states, which is capable of acquiring global information of the objective function. In contrast, for SA, if the objective function contains several high potential barriers, it may fall into local optima, as the thermal transition depends on the height of the potential barriers. To solve the problem, quantum tunneling is used to go through potential barriers. Figure 5 illustrates quantum tunneling versus simulated annealing, as done in Wang et al. (2024). A comparison between QAA and SA (Kadowaki and Nishimori 1998; Santoro et al. 2002; Battaglia et al. 2005; Crosson and Harrow 2016) suggests that QAA can be exponentially faster than SA in some cases.

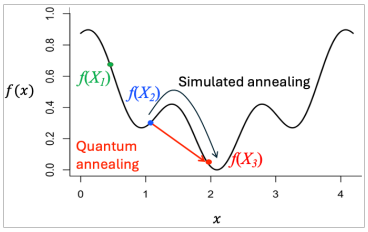

Figure 5: Comparison of the working principle of simulated annealing and QAA, as in Wang et al. (2024).

To apply quantum computing to optimization, an efficient encoding of variable and constraint spaces is necessary. This involves an intricate mapping of the optimization problem to the available number of qubits while preserving problem structure, which is challenging. Moreover, a challenge to quantum computing is the limited coherence durations and the need to maintain fragile quantum states for the duration of the execution. While quantum error correction offers a theoretical pathway to mitigate decoherence effects, its practical implementation at scale remains a substantial engineering and algorithmic hurdle. Consequently, the development of quantum optimization algorithms that are both theoretically and practically sound within the operational limitations of near-term quantum hardware constitutes a crucial area of ongoing development.

6 CONCLUSION

In conclusion, optimizing complex systems can be viewed in many different ways, depending on the user's overall goals. There are many methods available to assist in achieving these goals. Taking a comprehensive view of optimization allows users to take advantage of new methods for data collection, powerful computation, and engineering expertise to enhance the process of optimizing complex systems. We hope the research to attain the theoretical ideal of efficient optimization is in your future.

ACKNOWLEDGMENTS

This work has been funded in part by the U.S. National Science Foundation grant CMMI-2204872.

REFERENCES

Ahuja, R. K., Ö. Ergun, J. B. Orlin, and A. P. Punnen. 2002. "A Survey of Very Large-Scale Neighborhood Search Techniques". *Discrete Applied Mathematics* 123:75–102 https://doi.org/10.1080/24725854.2018.1490046.

Al-Dujaili, A., and S. Suresh. 2019. "Revisiting Norm Optimization for Multi-objective Black-box Problems: A Finite-time Analysis". *Journal of Global Optimization* 73(3):659–673 https://doi.org/10.1016/S0166-218X(01)00338-9.

Amaran, S., N. V. Sahinidis, B. Sharda, and S. J. Bury. 2016. "Simulation Optimization: A Review of Algorithms and Applications". *Annals of Operations Research* 240:351–380 https://doi.org/10.1007/s10479-015-2019-x.

Bajaj, I., A. Arora, and M. M. F. Hasan. 2021. "Black-Box Optimization: Methods and Applications". In *Black Box Optimization, Machine Learning, and No-Free Lunch Theorems*, edited by P. M. Pardalos, V. Rasskazova, and M. N. Vrahatis, 35–65. Cham: Springer International Publishing https://doi.org/https://doi.org/10.1007/978-3-030-66515-9_2.

Baritompa, W. P., D. W. Bulger, and G. R. Wood. 2005. "Grover's Quantum Algorithm Applied to Global Optimization". *SIAM Journal on Optimization* 15(4):1170–1184 https://doi.org/10.1137/040605072.

Barton, R. R. 2020. "Tutorial: Metamodeling for Simulation". In *2020 Winter Simulation Conference (WSC)*, 1102–1116 https://doi.org/10.1109/WSC48552.2020.9384059.

Battaglia, D., L. Stella, O. Zagordi, G. E Santoro, and E. Tosatti. 2005. "Deterministic and Stochastic Quantum Annealing Approaches". *Quantum Annealing and Other Optimization Methods* 679:171–206.

Boser, B. E., I. M. Guyon, and V. N. Vapnik. 1992. "A Training Algorithm for Optimal Margin Classifiers". In *COLT '92: Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. July 27-29, 1992, New York, 144–152. https://doi.org/10.1145/130385.130401.

Boyer, M., G. Brassard, P. Høyer, and A. Tapp. 1998. "Tight Bounds on Quantum Searching". *Fortschritte der Physik: Progress of Physics* 46(4-5):493–505.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 2017. *Classification and Regression Trees*. 1st ed. Oxfordshire: Routledge.

Bulger, D. W. 2007. "Combining a Local Search and Grover's Algorithm in Black-box Global Optimization". *Journal of Optimization Theory and Applications* 133:289–301 https://doi.org/10.1007/s10957-007-9168-2.

Bulger, D. W., W. P. Baritompa, and G. R. Wood. 2003. "Implementing Pure Adaptive Search with Grover's Quantum Algorithm". *Journal of Optimization Theory and Applications* 116(3):517–529.

Bulger, D. W., and G. R. Wood. 1998. "Hesitant Adaptive Search for Global Optimisation". *Mathematical Programming* 81:89–102.

Chen, C.-H., and L.-H. Lee. 2011. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. Singapore: World Scientific Publishing Co. Pte. Ltd.

Chen, E. J., and L. H. Lee. 2009. "A Multi-objective Selection Procedure of Determining a Pareto Set". *Computers & Operations Research* 36(6):1872–1879.

Crosson, E., and A. W. Harrow. 2016. "Simulated Quantum Annealing can be Exponentially Faster than Classical Simulated Annealing". In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. October 9-11, 2016, New York, 714–723.

De Haan, L. 1981. "Estimation of the Minimum of a Function Using Order Statistics". *Journal of the American Statistical Association* 76(374):467–469 https://doi.org/10.2307/228785.

Deb, K. 2001. *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester, New York: John Wiley & Sons.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II". *IEEE Transactions on Evolutionary Computation* 6:182–197 https://doi.org/10.1109/4235.996017.

Devroye, L., L. Gyorfi, A. Krzyzak, and G. Lugosi. 1994. "On the Strong Universal Consistency of Nearest Neighbor Regression Function Estimates". *The Annals of Statistics* 22(3):1371–1385.

Dürr, C. and P. Hoyer 1996. "A Quantum Algorithm for Finding the Minimum". arXiv preprint arXiv: 9607014.

Fu, M. C. 2015. *Handbook of Simulation Optimization*. New York: Springer.

Goodfellow, I., Y. Bengio, and A. Courville. 2016. *Deep Learning*. Cambridge: MIT Press.

Gosavi, A. 2015. *Simulation-Based Optimization*. 2nd ed. New York: Springer.

Graesser, D. L., Z. B. Zabinsky, M. E. Tuttle, and G. I. Kim. 1991. "Designing Laminated Composites Using Random Search Techniques". *Composite Structures* 18:311–325 https://doi.org/10.1016/0263-8223(91)90002-G.

Graesser, D. L., Z. B. Zabinsky, M. E. Tuttle, and G. I. Kim. 1993. "Optimal Design of a Composite Structure". *Composite Structures* 24:273–281.

Grover, L. K. 1996. "A Fast Quantum Mechanical Algorithm for Database Search". In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*. May 22-24, 1996, New York, 212-219.

Hansen, N., and A. Ostermeier. 2001. "Completely Derandomized Self-adaptation in Evolution Strategies". *Evolutionary Computation* 9(2):159–195.

Huang, H., and Z. B. Zabinsky. 2014. "Multiple Objective Probabilistic Branch and Bound for Pareto Optimal Approximation". In *2014 Winter Simulation Conference*, 3916–3927 https://doi.org/10.1109/WSC.2014.7020217.

Kadowaki, T., and H. Nishimori. 1998. "Quantum Annealing in the Transverse Ising Model". *Physical Review E* 58(5):5355.

Kiatsupaibul, S., R. L. Smith, and Z. B. Zabinsky. 2018. "Single Observation Adaptive Search for Continuous Simulation Optimization". *Operations Research* 66(6):1713–1727.

Kiatsupaibul, S., R. L. Smith, and Z. B. Zabinsky. 2020. "Single Observation Adaptive Search for Discrete and Continuous Stochastic Optimization". *Operations Research Letters* 48(5):666–673.

Kristinsdottir, B. P., Z. B. Zabinsky, M. E. Tuttle, and T. Csendes. 1996. "Incorporating Manufacturing Tolerances in Near-optimal Design of Composite Structures". *Engineering Optimization* 26:1–23.

Kristinsdottir, B. P., Z. B. Zabinsky, M. E. Tuttle, and S. Neogi. 2001. "Optimal Design of Large Composite Panels with Varying Loads". *Composite Structures* 51:93–102.

Küçükyavuz, S., and R. Jiang. 2022. "Chance-constrained Optimization Under Limited Distributional Information: A Review of Reformulations Based on Sampling and Distributional Robustness". *EURO Journal on Computational Optimization* 10:100030.

Li, H., L. H. Lee, E. P. Chew, and P. Lendermann. 2015. "MO-COMPASS: A Fast Convergent Search Algorithm for Multi-objective Discrete Optimization Via Simulation". *IIE Transactions* 47(11):1153–1169.

Li, J., W. Liu, G. Pedrielli, L. H. Lee, and E. P. Chew. 2017. "Optimal Computing Budget Allocation to Select the Nondominated Systems—A Large Deviations Perspective". *IEEE Transactions on Automatic Control* 63(9):2913–2927.

Li, K. and F. Li 2024. "Multi-fidelity Methods for Optimization: A Survey". arXiv preprint arXiv: 2402.09638.

Linz, D. D., Z. B. Zabinsky, S. Kiatsupaibul, and R. L. Smith. 2017. "A Computational Comparison of Simulation Optimization Methods Using Single Observations within a Shrinking Ball on Noisy Black-Box Functions with Mixed Integer and Continuous Domains". In *2017 Winter Simulation Conference (WSC)*, 2045–2056 https://doi.org/10.1109/WSC.2017.8247938.

Liu, Y., and G. J. Koehler. 2010. "Using Modifications to Grover's Search Algorithm for Quantum Global Optimization". *European Journal of Operational Research* 207(2):620–632.

Locatelli, M., and F. Schoen. 2021. "(Global) Optimization: Historical Notes and Recent Developments". *EURO Journal on Computational Optimization* 9:100012 https://doi.org/https://doi.org/10.1016/j.ejco.2021.100012.

Maneekul, P. and Z. B. Zabinsky and G. Pedrielli. 2025. "Branching Adaptive Surrogate Search Optimization (BASSO)". arXiv preprint arXiv: 2504.18002.

Mason, B., L. Jain, S. Mukherjee, R. Camilleri, K. Jamieson, and R. Nowak. 2022. "Nearly Optimal Algorithms for Level Set Estimation". In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*. March 28-30, 2022, 7625-7658.

Mete, H. O., and Z. B. Zabinsky. 2014. "Multiobjective Interacting Particle Algorithm for Global Optimization". *INFORMS Journal on Computing* 26(3):500–513.

Morey, D. F., G. Pedrielli, C. Wakayama, and Z. B. Zabinsky. 2024. "Multiple Model Consistency for Global Optimization". *Presentation at INFORMS Annual Meeting 2024, Seattle, WA*.

Morey, D. F., G. Pedrielli, and Z. B. Zabinsky. 2024. "A Hybrid Approach Combining Simulation and a Queueing Model for Optimizing a Biomanufacturing System". In *2024 Winter Simulation Conference (WSC)*, 1130–1138 https://doi.org/10.1109/WSC63780.2024.10838875.

Morey, D. F., R. Plate, C. Wakayama, and Z. B. Zabinsky. 2021. "Multi-fidelity Modeling for the Design of a Maritime Environmental Survey Network Utilizing Unmanned Underwater Vehicles". In *2021 Winter Simulation Conference (WSC)*, 1–12 https://doi.org/10.1109/WSC52266.2021.9715365.

Morimoto, K., Y. Takase, K. Mitarai, and K. Fujii. 2024. "Continuous Optimization by Quantum Adaptive Distribution Search". *Physical Review Research* 6(2):023191.

Peherstorfer, B., K. Willcox, and M. Gunzburger. 2018. "Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization". *SIAM Review* 60:550–591 https://doi.org/10.1137/16M1082469.

Ploskas, N., and N. V. Sahinidis. 2022. "Review and Comparison of Algorithms and Software for Mixed-integer Derivative-free Optimization". *Journal of Global Optimization* 82(3):433–462.

Protopopescu, V., and J. Barhen. 2005. "Quantum Algorithm for Continuous Global Optimization". In *Optimization and Control with Applications*, edited by L. Qi, K. Teo, and X. Yang, 293–303. Boston, MA: Springer US.

Rios, L. M., and N. V. Sahinidis. 2013. "Derivative-free Optimization: A Review of Algorithms and Comparison of Software Implementations". *Journal of Global Optimization* 56:1247–1293 https://doi.org/10.1007/s10898-012-9951-y.

Rockafellar, R. T., and S. Uryasev. 2000. "Optimization of Conditional Value-at-risk". *Journal of Risk* 2:21–42.

Rojas-Gonzalez, S., and I. Van Nieuwenhuyse. 2020. "A Survey on Kriging-based Infill Algorithms for Multiobjective Simulation Optimization". *Computers & Operations Research* 116:104869.

Romeijn, H. E., and R. L. Smith. 1994. "Simulated Annealing and Adaptive Search in Global Optimization". *Probability in the Engineering and Informational Sciences* 8:571–590.

Santoro, G. E., R. Martonák, E. Tosatti, and R. Car. 2002. "Theory of Quantum Annealing of an Ising Spin Glass". *Science* 295(5564):2427–2430.

Sarker, R., K.-H. Liang, and C. Newton. 2002. "A New Multiobjective Evolutionary Algorithm". *European Journal of Operational Research* 140(1):12–23.

Savic, V., M. E. Tuttle, and Z. B. Zabinsky. 2001. "Optimization of Composite I-Sections Using Fiber Angles as Design Variables". *Composite Structures* 53:265–277.

Shashaani, S. 2024. "Simulation Optimization: An Introductory Tutorial on Methodology". In *2024 Winter Simulation Conference (WSC)*, 1308–1322 https://doi.org/10.1109/WSC63780.2024.10838809.

Shen, Y., S. Kiatsupaibul, Z. B. Zabinsky, and R. L. Smith. 2007. "An Analytically Derived Cooling Schedule for Simulated Annealing". *Journal of Global Optimization* 38(3):333–365 https://doi.org/10.1007/s10898-006-9068-2.

Shi, L., and S. Ólafsson. 2009. *Nested Partitions Method, Theory and Applications*. New York: Springer.

Shor, P. W. 1994. "Algorithms for Quantum Computation: Discrete Logarithms and Factoring". In *Proceedings 35th Annual Symposium on Foundations of Computer Science*. November 20-22, 1994, Santa Fe, NM, 124–134.

Smith, R. L. 1984. "Efficient Monte Carlo Procedures for Generating Points Uniformly Distributed Over Bounded Regions". *Operations Research* 32(6):1296–1308.

Stephens, C., and W. Baritompa. 1998. "Global Optimization Requires Global Information". *Journal of Optimization Theory and Applications* 96:575–588.

Törn, A., and A. Žilinskas. 1989. *Global Optimization*. Heidelberg: Springer.

Uryasev, S., and R. T. Rockafellar. 2001. "Conditional Value-at-Risk: Optimization Approach". In *Stochastic Optimization: Algorithms and Applications*, edited by S. Uryasev and P. M. Pardalos, 411–435. Boston, MA: Springer US https://doi.org/10.1007/978-1-4757-6594-6_17.

Verma, S., M. Pant, and V. Snasel. 2021. "A Comprehensive Review on NSGA-II for Multi-objective Combinatorial Optimization Problems". *IEEE Access* 9:57757–57791.

Vu, K. K., C. d'Ambrosio, Y. Hamadi, and L. Liberti. 2017. "Surrogate-based Methods for Black-box Optimization". *International Transactions in Operational Research* 24(3):393–424.

Wang, X., Z. Wang, and B. Ni. 2024. "Mapping Structural Topology Optimization Problems to Quantum Annealing". *Structural and Multidisciplinary Optimization* 67(5):74.

Wilson, J., F. Hutter, and M. Deisenroth. 2018. "Maximizing Acquisition Functions for Bayesian Optimization". *Advances in Neural Information Processing Systems* 31:9906–9917.

Wood, G. R., and Z. B. Zabinsky. 2002. "Stochastic Adaptive Search". In *Handbook of Global Optimization*, edited by P. M. Pardalos and H. E. Romeijn, Volume 2, 231–249. Boston, MA: Springer US https://doi.org/10.1007/978-1-4757-5362-2_7.

Wood, G. R., Z. B. Zabinsky, and B. P. Kristinsdottir. 2001. "Hesitant Adaptive Search: The Distribution of the Number of Iterations to Convergence". *Mathematical Programming* 89(3):479–486.

Xu, J., S. Zhang, E. Huang, C. H. Chen, L. H. Lee, and N. Celik. 2016. "MO2TOS: Multi-Fidelity Optimization with Ordinal Transformation and Optimal Sampling". *Asia-Pacific Journal of Operational Research* 33:3737 – 3748 https://doi.org/10.1142/S0217595916500172.

Yang, R., L. Wang, and Z. Wang. 2011. "Multi-objective Particle Swarm Optimization for Decision-making in Building Automation". In *2011 IEEE Power and Energy Society General Meeting*. July, 24-28, 2011, Detroit, MI, 1-5.

Yoon, M., and J. Bekker. 2020. "Multi-objective Simulation Optimisation on Discrete Sets: A Literature Review". *International Journal of Operational Research* 39(3):364–405.

Yusoff, Y., M. S. Ngadiman, and A. M. Zain. 2011. "Overview of NSGA-II for Optimizing Machining Process Parameters". *Procedia Engineering* 15:3978–3983.

Zabinsky, Z. B. 2003. *Stochastic Adaptive Search for Global Optimization*. New York: Springer https://doi.org/doi.org/10.1007/978-1-4419-9182-9.

Zabinsky, Z. B., and H. Huang. 2020. "A Partition-based Optimization Approach for Level Set Approximation: Probabilistic Branch and Bound". In *Women in Industrial and Systems Engineering: Key Advances and Perspectives on Emerging Topics*, edited by A. E. Smith, 113–155. Cham: Springer Nature.

Zabinsky, Z. B., and D. D. Linz. 2023. "Hesitant Adaptive Search with Estimation and Quantile Adaptive Search for Global Optimization with Noise". *Journal of Global Optimization* 87:31–55.

Zabinsky, Z. B., G. Pedrielli, and H. Huang. 2019. "A Framework for Multi-Fidelity Modeling in Global Optimization Approaches". In *Proceedings of Fifth International Conference on Machine Learning, Optimization, and Data Science (LOD2019)*. September 10-13, 2019, Siena, Italy.

Zabinsky, Z. B., and R. L. Smith. 1992. "Pure Adaptive Search in Global Optimization". *Mathematical Programming* 53(1-3):323–338 https://doi.org/10.1007/BF01585710.

Zabinsky, Z. B., R. L. Smith, J. F. McDonald, H. E. Romeijn, and D. E. Kaufman. 1993. "Improving Hit-and-Run for Global Optimization". *Journal of Global Optimization* 3(2):171–192.

Zabinsky, Z. B., M. E. Tuttle, and C. Khompatraporn. 2006. "A Case Study: Composite Structure Design Optimization". In *Global Optimization: Scientific and Engineering Case Studies*, edited by J. Pintér, 507–528. New York: Springer Science and Business Media.

Zabinsky, Z. B., G. R. Wood, M. A. Steel, and W. P. Baritompa. 1995. "Pure Adaptive Search for Finite Global Optimization". *Mathematical Programming* 69(1-3):443–448.

AUTHOR BIOGRAPHIES

ZELDA B. ZABINSKY is an Emeritus Professor in Industrial and Systems Engineering at the University of Washington in Seattle, WA. She is an INFORMS Fellow and an IISE Fellow. Her Ph.D. is from the University of Michigan, in Industrial and Operations Engineering. Her research interests are in global optimization under uncertainty for complex systems, and she has worked in many application areas. Her email address is zelda@uw.edu. Her website is https://ise.washington.edu/facultyfinder/zelda-zabinsky.

PARIYAKORN MANEEKUL is a Ph.D. candidate in the Department of Industrial and Systems Engineering at the University of Washington. Her research interest includes optimization under uncertainty and the interplay of statistical learning methods and optimization. Her email address is parim@uw.edu.